CLAIMS:

1.   A method of summarizing digital audio data comprising the steps of:

   directly analyzing the audio data to identify a representation of the audio data having at least one calculated feature characteristic of the audio data;

   classifying the audio data on the basis of the representation into a category selected from at least two categories; and

   generating an acoustic signal representative of a summarization of the digital audio data, wherein the summarization is dependent on the selected category.

2.   A method as claimed in claim 1, wherein the analyzing step further comprises segmenting audio data into segment frames, and overlapping the frames.

3.   A method as claimed in claim 2, wherein the classifying step further comprises classifying the frames into a category by collecting training data from each frame and determining classification parameters by using a training calculation.

4.   A method as claimed in any preceding claim, wherein the calculated feature comprises perceptual and subjective features related to music content.

5.   A method as claimed in claim 3, wherein the training calculation comprises a statistical learning algorithm

wherein the statistical learning algorithm is Hidden Markov Model, Neural Network, or Support Vector Machine.

6. A method as claimed in any preceding claim, wherein the type of acoustic signal is music.

7. A method as claimed in any preceding claim, wherein the type of acoustic signal is vocal music or pure music.

8. A method as claimed in any preceding claim, wherein the calculated feature is amplitude envelope, power spectrum or mel-frequency cepstral coefficients.

9. A method as claimed in any preceding claim, wherein the summarization is generated in terms of clustered results and heuristic rules related to pure or vocal music.

10. A method as claimed in any preceding claim, wherein the calculated feature relates to pure or vocal music content and is linear prediction coefficients, zero crossing rates, or mel-frequency cepstral coefficients.

11. An apparatus for summarizing digital audio data comprising:

a feature extractor for receiving audio data and directly analyzing the audio data to identify a representation of the audio data having at least one calculated feature characteristic of the audio data;

a classifier in communication with the feature extractor for classifying the audio data on the basis of the

ART 34 AMDT

representation received from the feature extractor into a category selected from at least two categories; and

a summarizer in communication with the classifier. for generating an acoustic signal representative of a summarization of the digital audio data, wherein the summarization is dependent on the category selected by the classifier.

12. An apparatus as claimed in claim 11, further comprising a segmentor in communication with the feature extractor for receiving an audio file and segmenting audio data into segment frames, and overlapping the frames for the feature extractor.

13. An apparatus as claimed in claim 12, further comprising a classification parameter generator in communication with the classifier, wherein the classifier classifies each of the frames into a category by collecting training data from each frame and determining classification parameters by using a training calculation in the classification parameter generator.

14. An apparatus as claimed in any of claims 11-13, wherein the calculated feature comprises perceptual and subjective features related to music content.

15. An apparatus as claimed in any of claims 11-14, wherein the training calculation comprises a statistical learning algorithm wherein the statistical learning algorithm is Hidden Markov Model, Neural Network, or Support Vector Machine.

16. An apparatus as claimed in any of claims 11-15, wherein the acoustic signal is music.

17. An apparatus as claimed in any of claims 11-16, wherein the acoustic signal is vocal music or pure music.

18. An apparatus as claimed in any of claims 11-17, wherein the calculated feature is amplitude envelope, power spectrum or mel-frequency cepstral coefficients.

19. An apparatus as claimed in any of claims 11-18, wherein the summarizer generates the summarization in terms of clustered results and heuristic rules related to pure or vocal music.

20. An apparatus as claimed in any of claims 11-19, wherein the calculated feature relates to pure or vocal music content and is linear prediction coefficients, zero crossing rates, or mel-frequency.

21. A computer program product for summarizing digital audio data comprising a computer usable medium having computer readable program code means embodied in said medium for causing the summarizing of digital audio data, said computer program product comprising:
    a computer readable program code means for directly analyzing the audio data to identify a representation of the audio data having at least one calculated feature characteristic of the audio data;

a computer readable program code for classifying the audio data on the basis of the representation into a category selected from at least two categories; and

a computer readable program code for generating an acoustic signal representative of a summarization of the digital audio data, wherein the summarization is dependent on the selected category.

22. A computer program product as claimed in claim 21, wherein analyzing further comprises segmenting audio data into segment frames, and overlapping the frames.

23. A computer program product as claimed in claim 22, wherein classifying further comprises classifying the frames into a category by collecting training data from each frame and determining classification parameters by using a training calculation.

24. A computer program product as claimed in any of claims 21-23, wherein the calculated feature comprises perceptual and subjective features related to music content.

25. A computer program product as claimed in any of claims 21-24, wherein the training calculation comprises a statistical learning algorithm wherein the statistical learning algorithm is Hidden Markov Model, Neural Network, or Support Vector Machine.

26. A computer program product as claimed in any of claims 21-25, wherein the acoustic signal is music.

27. A computer program product as claimed in any of claims 21-26, wherein the type of acoustic signal is vocal music or pure music.

28. A computer program product as claimed in any of claims 21-27, wherein the calculated feature is amplitude envelope, power spectrum or mel-frequency cepstral coefficients.

29. A computer program product as claimed in any of claims 21-28, wherein the summarization is generated in terms of clustered results and heuristic rules related to pure or vocal music.

30. A computer program product as claimed in any of claims 21-29, wherein the calculated feature relates to pure or vocal music content and is linear prediction coefficients, zero crossing rates, or mel-frequency.